

CRIMES IN ATLANTA

ANDRES SOTELO

OBJECTIVE

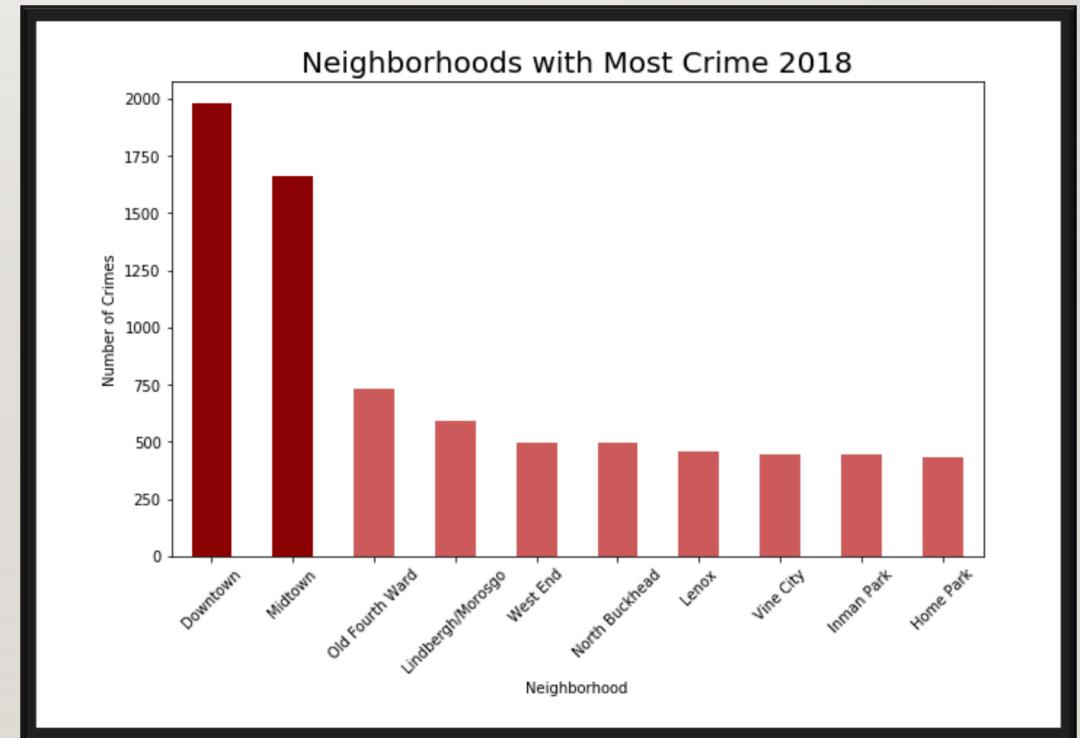
The objective of this analysis is to find a correlation of venues and crime rate of different neighborhoods in Atlanta. Venues typically consists of parks, restaurants/bars, museums and much more. If we find a correlation between certain types of venues and crime, then the city of Atlanta can use that information to take necessary action. I will also cluster neighborhoods which will discover similarities of those neighborhoods in regards to their respective crime rate.

DATA

- There are two sources from where I gathered the data:
 - https://en.wikipedia.org/wiki/Neighborhood_planning_unit
 - <opendata.atlantapd.org/Crimedata/Default.aspx>
- The first link is a list of the neighborhoods in Atlanta and the second is crime data from Atlanta PD which includes over 30,000 observation over the span of January 1, 2019 through December 31, 2019.
- It is very important to mention that I had to remove missing rows, rows that had the same information (duplicates), remove unnecessary columns and more to clean the data.

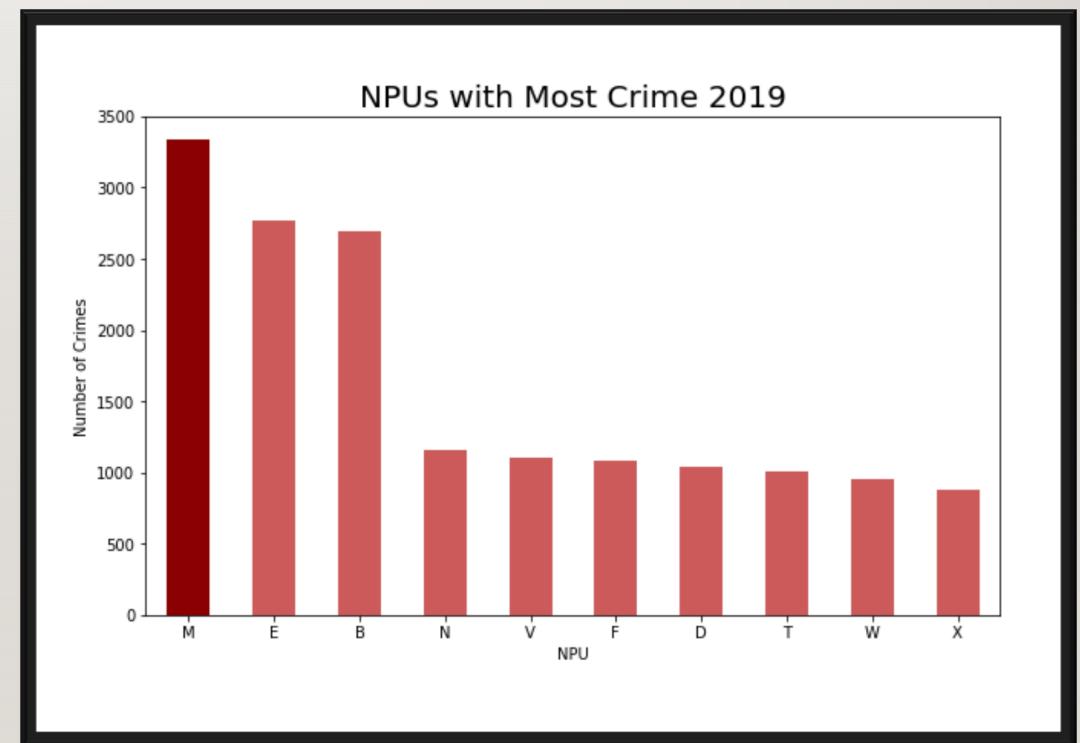
DATA EXPLORATION

- This chart represents the neighborhoods with higher crime rate. Based on the data, Downtown and Midtown are the top neighborhoods with the highest crime rate overall. (By the way, it should say “Neighborhoods with Most Crime 2019”)

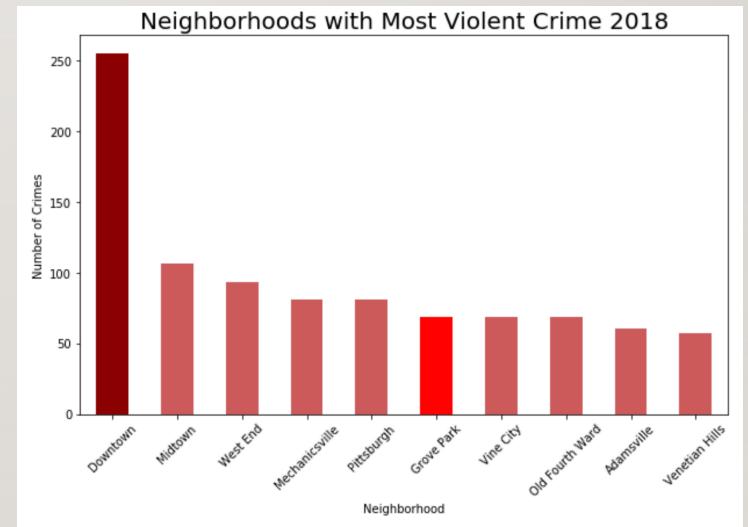
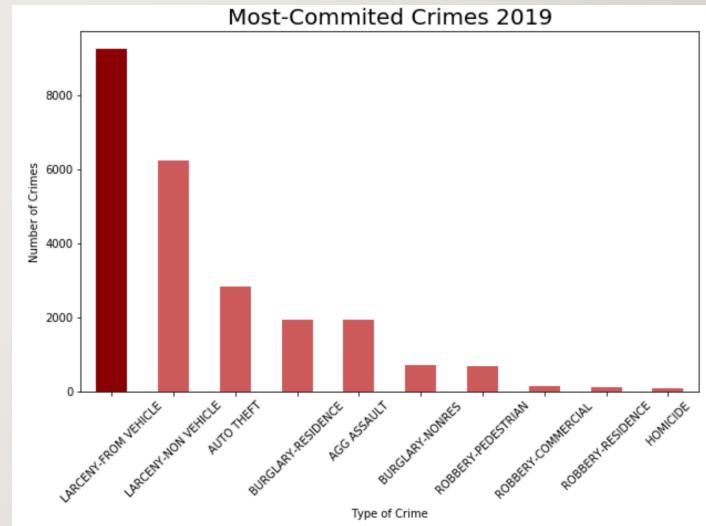


DATA EXPLORATION (CONT'D)

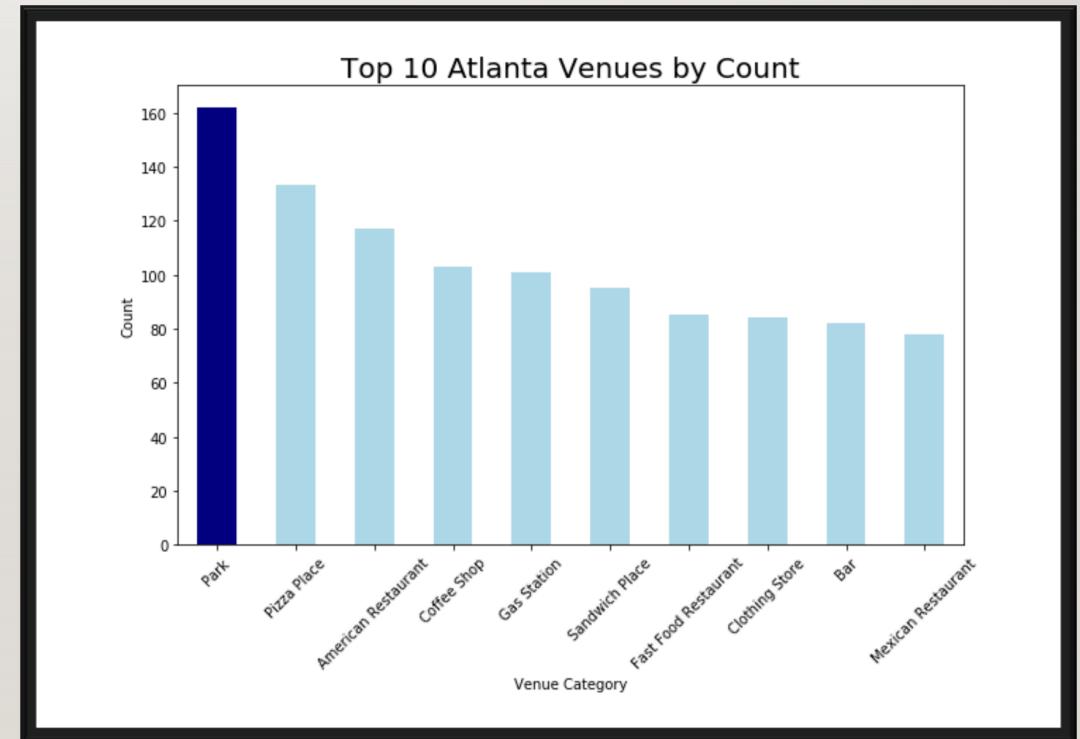
- This chart represent the NPU's with the higher crime rates. Remember, NPU stands for Neighborhood Planning Unit.



- The top chart categorizes the most committed crime in 2019 and we see that Larceny takes 1st and 2nd place.
- The bottom chart represents the neighborhoods with the highest violent crime rate. According to the data, Downtown and Midtown have the highest number of violent crime.

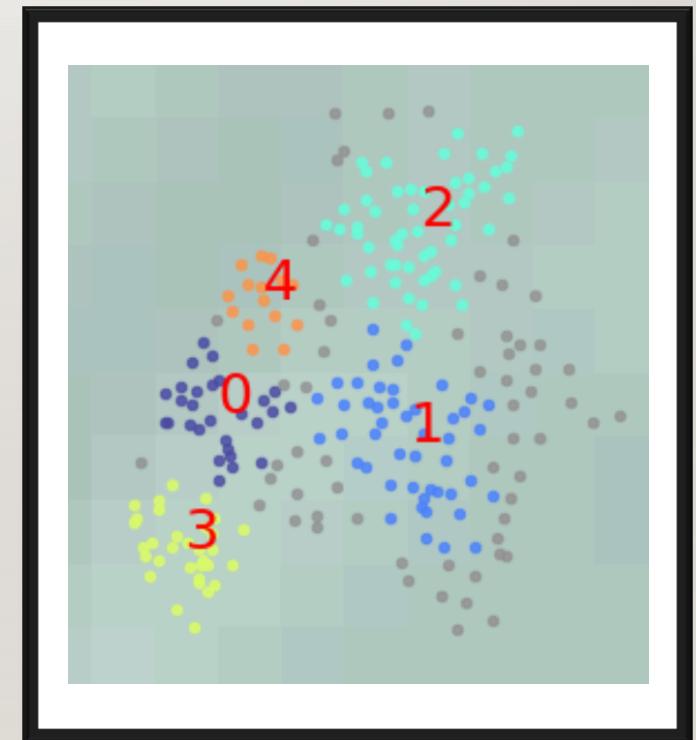


-
- As we can see from our bar chart, the top venues are parks, pizza and American restaurants in 2019.



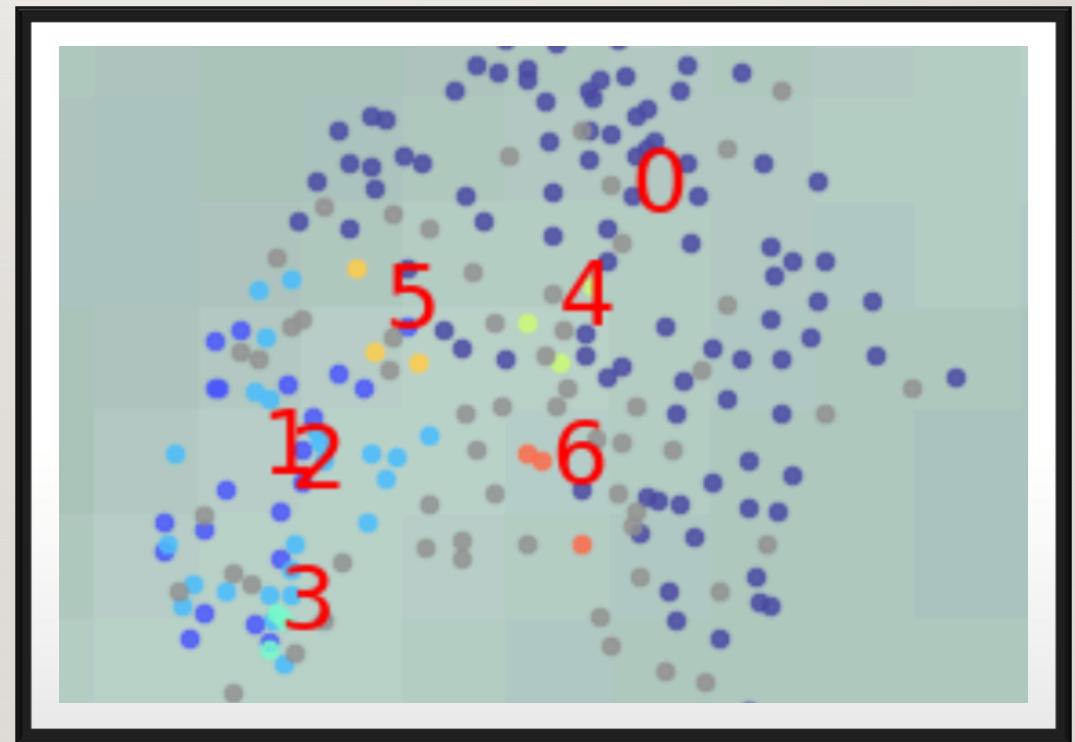
DBSCAN CLUSTERING BY LOCATION

- The visualization represents how many neighborhoods we can cluster based on location. The number of average crimes, overall in each cluster are:
 - Cluster 0: 34
 - Cluster 1: 172
 - Cluster 2: 91
 - Cluster 3: 33
 - Cluster 4: 47



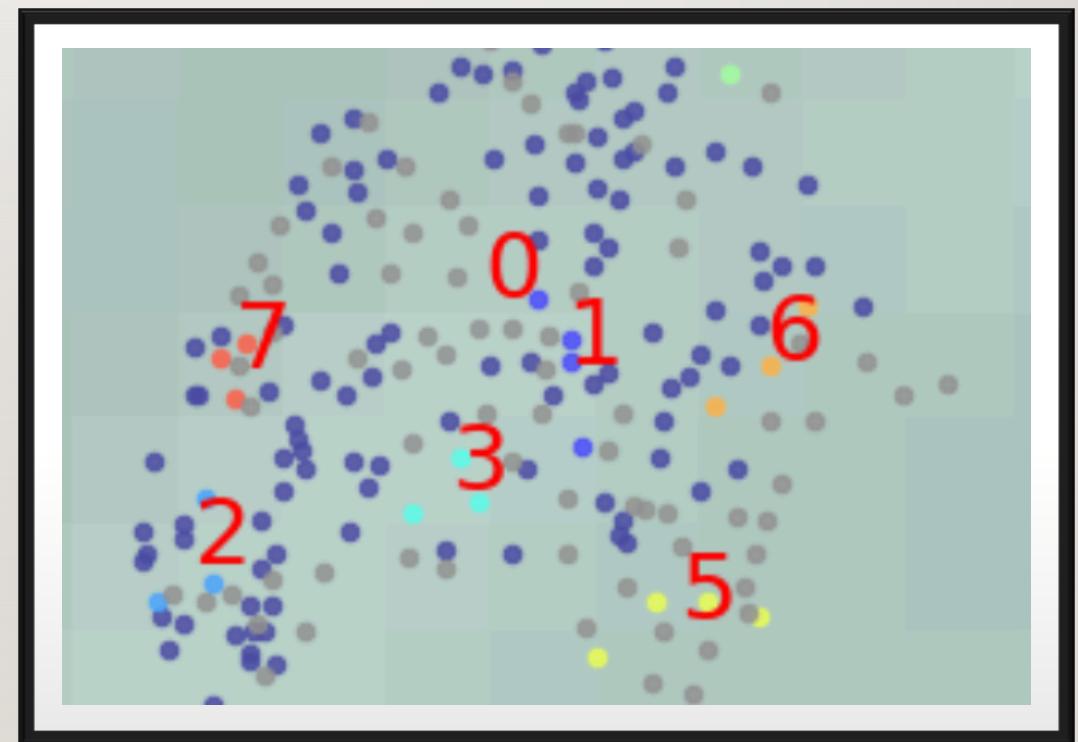
DBSCAN CLUSTER BY VIOLENT CRIMES

- This visualization represents a cluster of neighborhoods based on violent crimes (assault and homicide).
 - Cluster 0: 130
 - Cluster 1: 7
 - Cluster 2: 15
 - Cluster 3: 6
 - Cluster 4: 151
 - Cluster 5: 154
 - Cluster 6: 219



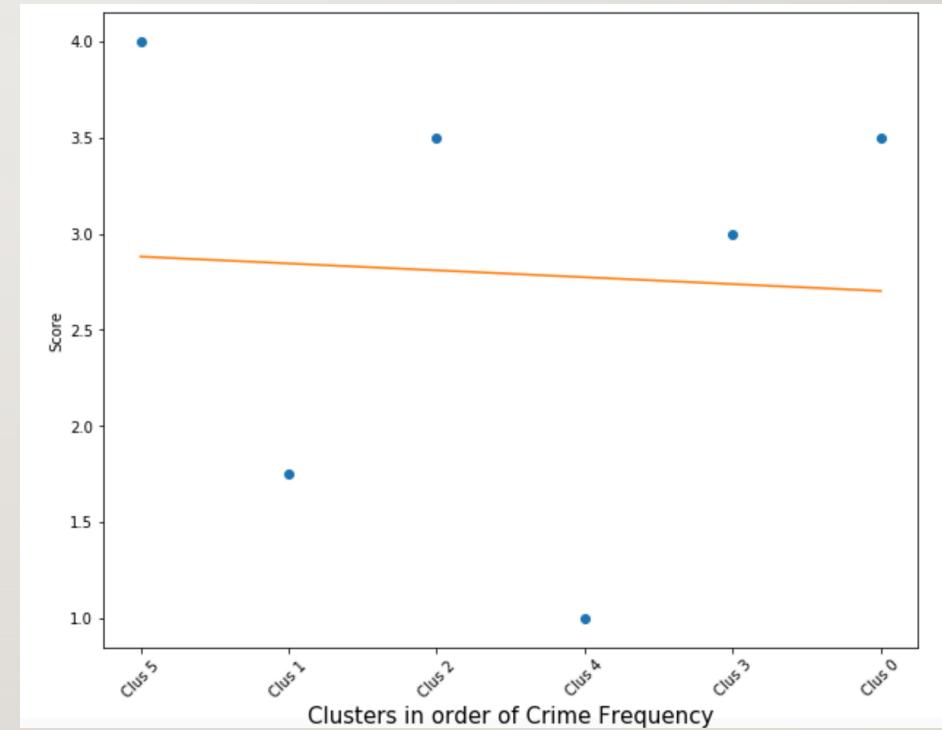
DBSCAN CLUSTERING BY ROBBERIES

- This visualization represents a cluster of neighborhoods based on robberies.
 - Cluster 0: 70
 - Cluster 1: 205
 - Cluster 2: 25
 - Cluster 3: 134
 - Cluster 4: 290
 - Cluster 5: 57
 - Cluster 6: 237
 - Cluster 7: 17



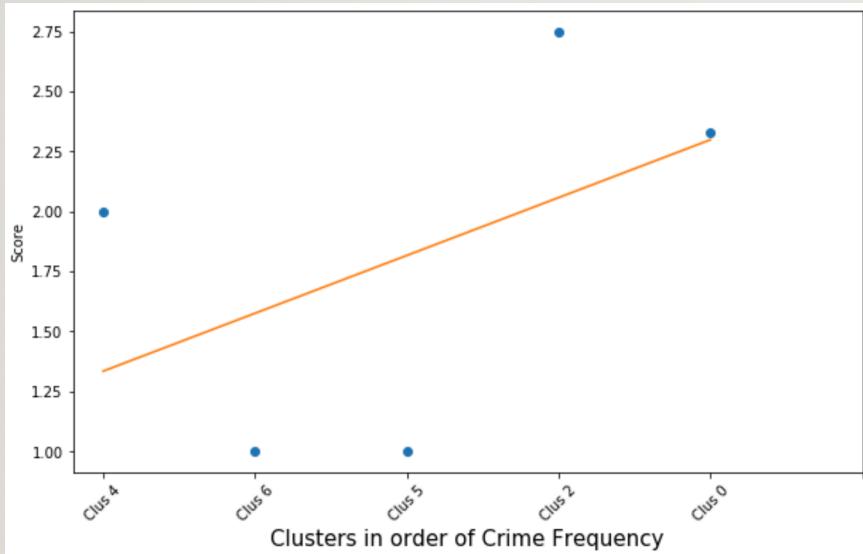
RESULTS

- The correlation between crime (overall crime) and location is slightly negative.

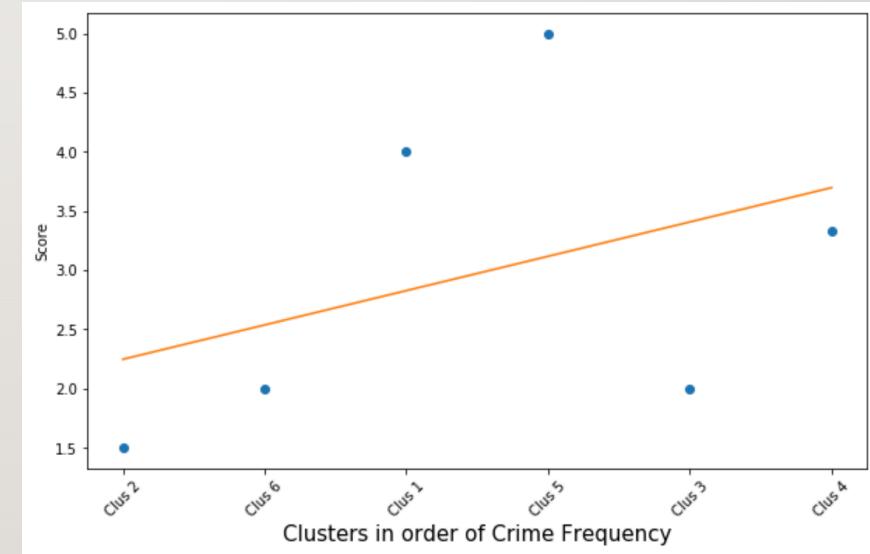


RESULTS

VIOLENT CRIMES VS NEIGHBORHOOD



ROBBERIES VS NEIGHBORHOOD



CONCLUSION

The venues I specifically chose to determine a correlation between crime (violent and robberies) and location are liquor stores and gun/pawn shops. As we can see, from the two previous scatterplots, there is a positive correlation. However, it is not guaranteed that if those venues were not to exist, then crime will go down. There may be other variables that need to be considered.